UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2016/37

Job flows, worker flows, and churning in South Africa

Andrew Kerr*

April 2016

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: In this paper, worker and job flows are estimated using the IRP5 data from the South African Revenue Services. The data used in this paper is from the 2011–14 tax years and contains information on more than 12 million individuals and nearly 300,000 firms. The main finding of the paper is that worker flows are substantial, more than 50 per cent per year. Job flows are higher than previous estimates. The findings suggest that rigidity in the labour market does not come from rigidities in employment adjustment.

Keywords: churning, job and worker flows, labour regulation
JEL classification: J21, J23, J63

Acknowledgements: The author wishes to thank the South African Revenue Service (SARS), UNU-WIDER, and the South African National Treasury for support. The opinions expressed in this paper do not reflect the views of SARS or the National Treasury. Without implicating any of them in the conclusions drawn in this paper, I thank Liz Gavin, Kholofelo Makua, and Michelle Smit from SARS; Aroop Chatterjee, Chris Axelson, Wiaan Boonzaaier, Marle Van Niekerk, and Landon McMillan from the National Treasury; as well as Channing Arndt, Lawrence Edwards, Friedrich Krueser, Tasha Naughtin, Vimal Ranchhod, Neil Rankin, Nicola Viegi, and Martin Wittenberg for helpful comments, suggestions, and advice on using the data. All output used in the paper has been checked by the National Treasury to ensure that it does not compromise the anonymity of any firm or worker.

* DataFirst, University of Cape Town, South Africa, Andrew.kerr@uct.ac.za.

This study has been prepared within the UNU-WIDER project on 'Firm- and Industry-level Analysis in South Africa', which is part of a larger research project on 'Regional growth and development in Southern Africa'.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-080-5

Typescript prepared by Sophie Richmond.

UNU-WIDER acknowledges specific programme contribution from the National Treasury of South Africa to its project 'Regional Growth and Development in Southern Africa' and core financial support to its work programme from the governments of Denmark, Finland, Sweden, and the United Kingdom

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Worker and job flows are a pervasive feature of all labour markets. Every year a substantial fraction of the workforce either start working for a new employer or leave an old employer. Simultaneously firms are contracting or expanding the number of workers they employ. Patterns of hiring and firing are complex, with shrinking firms hiring and growing firms firing workers (Hamermesh et al. 1996). Measuring and describing these flows is a fairly recent phenomenon, however, even in advanced economies, since the data requirements are extensive. In this paper the extent of job flows, worker flows, and churning in South Africa is documented using a new source of administrative data derived from anonymized tax records. A better understanding of job and worker flows is important in South Africa because these shed light on labour demand, labour reallocation and the effects of the regulatory environment on the labour market, which are not well understood and which are critical if South Africa's extremely high rate of unemployment is to be lowered.

New data made available by the South African Revenue Service (SARS) makes the measurement and analysis of job and worker flows possible in South Africa. These data are a census of all workers earning more than ZAR2,000 per year in Pay As You Earn (PAYE) tax registered firms, containing some basic individual information, including individual identifiers, together with firm identifiers and some limited firm information. The data can be made into a panel of workers and then also of firms. This firm dataset is comprised of tax-paying, and thus formal, firms. The exclusion of firms not registered for PAYE tax is not as much of an issue in South Africa as in other developing countries because South Africa has a smaller informal economy relative to these countries (Magruder 2012). The total number of individuals in the data employed in March 2014 is approximately 9.3 million compared to the estimate from the South African Quarterly Labour Force Survey of approximately 15 million from the first quarter of 2014. The tax data thus cover approximately 62 per cent of all employment in the country. Any worker not working in a PAYE registered firm would be excluded from the IRP5 data.

In this paper the extent of worker and job flows is estimated using the SARS IRP5 data. Worker flows are found to be substantial, higher than all but two of the worker flow rates estimated for 24 OECD (Organisation for Economic Co-operation and Development) countries by Bassanini et al. (2010).[1] This is perhaps surprising given the reputation of the South African labour market as rigid (see Go et al. 2009). The finding in this paper of extensive worker flows when using the IRP5 tax data is consistent with previous work using worker-level panel data by Banerjee et al. (2008), and contradicts the assertion of Go et al. (2009) that the high level of worker flows estimated by Banerjee et al. (2008) was due to more fluid transitions between non-employment and the informal sector. A key finding of this paper is that even in the SARS IRP5 data of individuals working for PAYE tax registered firms there are still very substantial worker flows.

In the following section the relevant literature is reviewed. Section 3 describes the SARS data used in the paper and the challenges faced in using it. In Section 4 the estimation of job and worker flows as well as churning is undertaken. Section 5 concludes.

[1] It should be noted that the Bassanini et al. (2010) worker flow estimates came from household survey data, but other than this the same method was used as in this paper.

2 Literature review

This section begins by defining key terms used in the job and worker flows literature. The different measuring concepts used to actualize the measurement of worker flows are then outlined and then the literature on job flows, worker flows, and churning is discussed.

2.1 Concepts

Job flows measure the gross creation and destruction of jobs, or the increasing or decreasing size of firms:

$$JF_{it} = E_{it}-E_{it-1} = H_{it} - S_{it}$$

where JF is job flows, E is employment, H is hires, and S is separations, all measured at firm i at a specific point in time. Gross job creation is the total change in employment across all firms that grew while gross job destruction is the total change in employment in all firms that shrank. Job reallocation is defined as the absolute value of job flows, and job creation in a firm is a positive job flow while job destruction in a firm is a negative job flow.

Worker flows are defined as the sum of hires and separations:

$$WF_{it} = H_{it}+S_{it}.$$

Job reallocation (JR) is the absolute value of job flows. It is then possible to link job flows, worker flows, and job reallocation so that:

$$WF_{it} = JR_{it} + CF_{it},$$

where CF_{it} is excess worker flows or churning, a residual component of worker flows above those resulting from job reallocation.

Following much of the literature, job flows, job reallocation, worker flows, and churning flows are summed across firms and then expressed as percentages of firm-level employment. Following Davis et al. (1996), the average of firm employment at *t* and *t*-1 is used to calculate rates in the rest of the paper.

2.2 Measurement of worker flows

Davis and Haltiwanger (1998) have noted that two different measuring concepts, reallocation and turnover, have been applied in the literature to actualize the measurement of worker flows, as defined above. The choice one or the other can mean substantial differences in measured worker flows (Huber and Smeral 2006) and may hamper comparisons across countries. Using the reallocation concept, worker flows are measured as 'the number of persons whose place of employment differs between *t*-1 and *t*' (Huber and Smeral 2006: 1689). Using the turnover conceptualization, worker flows are measured as 'the number of accessions plus the number of separations that occur from *t*-1 to [*t*].' The key difference between the two measuring concepts is how short-term spells are treated. Using the turnover concept, short-term spells that both start and end between *t*-1 and *t* are not counted. Such short-term spells are counted using the reallocation concept however.

Huber and Smeral (2006) have argued that there are important practical differences between the two ways of measuring worker flows, but that little attention has been paid to this issue. The

SARS IRP5 data potentially allows for the use of both measuring concepts because on each tax certificate issued a firm should record the date employed from and employed to in the tax year. However, due to measurement error concerns about these employment dates, in this paper a turnover measure is used, asking how many workers have changed employers between one week in a particular year and the same week in the next year. Worker flows will thus be under-counted by missing short-term spells of employment that start and end within this year.

2.3 The extent of worker flows

A detailed picture of worker flows and churning in the United States (US) using firm data emerged from the pioneering work of Burgess et al. (2000), who analysed quarterly administrative data on workers and firms from the state of Maryland.[2] This work showed that worker reallocation over and above job reallocation, or churning, was very high in the US, and in fact was larger than job reallocation. Churning was also found to be important in all industries. Earlier work had shown that firm-level job flows were very heterogeneous even within narrowly defined industries (Davis et al. 1996). Burgess et al. (2000) showed this was true for churning flows but that employer fixed effects were much more important in explaining churning flows than they were in explaining job flows. They also showed that there was substantial persistence in employer level churning rates. These facts led the authors to suggest that different human resource policies were used by different firms even within narrowly defined industries. These issues are taken up in the analysis of the SARS IRP5 data below.

Subsequently there have been a number of studies showing that worker flows are high in a variety of countries and a few studies comparing worker flows across countries. The cross-country worker flows studies generally use worker data from household panel surveys without firm identifiers (see Bassanini et al. 2010; OECD 2009), due to the difficulty of obtaining and analysing comparable administrative data on workers and their firms across countries. Bassanini et al. (2010) use worker reallocation measures to study worker flows in 25 mostly OECD countries. The authors show that while worker reallocation rates are substantial, there is also a lot of variation in yearly worker reallocation rates, which range from 26 per cent in Greece to nearly 50 per cent in the US and 56 per cent in Iceland, with the median countries, France and Switzerland, having reallocation rates of 37.7 per cent and 35.9 per cent respectively.

2.4 Worker flows and labour market rigidities

In a review of the literature on the relationship between worker flows and labour market rigidities, Bassanini et al. (2010) argue that higher worker flows are a result of lower rigidities in the labour market. Employment protection, unemployment benefits, minimum wages, and product market regulation may all affect worker flows. The authors argue that employment protection in the form of firing restrictions or limiting the use of temporary employment contracts is assumed to reduce worker and job flows.

Martin and Scarpetta (2012) provide a more nuanced view. They argue that if some contracts or types of employment are better protected and others less so this could lead to an over-reliance on the less protected contracts and give rise to higher worker flows. This over-reliance would be less efficient since the authors argue it could lead to less training and investment in firm–worker matches by both parties.

---

[2] There is a much older literature looking at worker flows, using data on individuals from household surveys, which is discussed below in the context of cross-country differences in worker flows.

To anticipate the results below, using the IRP5 data South Africa has levels of worker flows above almost all of the 25 OECD countries for which Bassanini et al. (2010) and OECD (2009) have data.[3] In the interpretation of Bassanini et al. (2010), this would then indicate low levels of employment protection, unemployment benefits, minimum wages, or product market regulation. But if one believes Martin and Scarpetta (2012), then this result may mean the existence of a dual labour market in which some workers have highly protected jobs while others do not. In this interpretation, the finding of high worker flows could still mean employment in some parts of the labour market is over-protected. The analysis below does provide some suggestion that, as in the US, frequent worker flows for some workers coexist with stable jobs for a sizeable fraction of workers, but it is important to note that the predominant view is that high worker flows are an indication that rigidities affecting firms' decisions to adjust their size are not substantial.

Dadam and Viegi (2015) provide evidence that, despite a relatively mild recession in 2009 compared to the United Kingdom and the US, South Africa experienced a massive loss of employment in the recession. This is consistent with evidence that South African employers can adjust employment and it is wages that are rigid, so that most of the adjustment in a recession is through employers decreasing employment rather than wages. It is also consistent with the findings below, that worker flows and churning are substantial in South Africa, indicating that rigidities around hiring and firing are not as substantial as has been argued (see Go et al. 2009).

2.5 Churning

Part of the interest in worker flows is that they have been found to be substantially above job flows, despite earlier labour market models assuming them to be equal (see Mortensen and Pissarides 1994). This means that there is substantially more movement of workers than the amount of movement needed by firms in their adjustment to their preferred sizes. As discussed above, the difference between job flows and worker flows is labelled churning. Burgess et al. (2000: 479) interpret churning as 'the reevaluation of a job match, initiated either by an employer, and evidenced by simultaneous hiring and firing, or by an employee, and evidenced by replacement of quits'. In this paper the importance of churning is measured relative to job flows, and the persistence of churning rates at the firm level is explored.

2.6 Job flows

Job flows measure the extent of changes in firm size. They reflect changes in firms' labour demand (Davis et al. 1996). Job flows have been studied in detail using firm-level panel data for nearly 30 years, and there is substantially more cross-country analysis of job flows than worker flows (see Criscuolo et al. 2014; Haltiwanger et al. 2008). The first and only paper looking at job flows using firm data in South Africa is Kerr et al. (2014), who use the Quarterly Employment Statistics (QES) firm survey undertaken by Statistics South Africa to look at job creation and destruction in South African firms. The main result is that levels of job creation and destruction are not substantially below OECD country levels, particularly because of some limitations with the data.

The authors noted two important limitations of the Statistics South Africa QES survey data they used. The first of these was that births were not well captured in the QES data. This was due to the nature of the QES panel—once a sample was selected no new firms entered the sample and the same sample was used for up to four years. The second limitation was that small firm job

[3] Again it should be noted that these estimates are constructed using household survey data rather than tax data as is used in this paper.

reallocation was not measured whenever a new sample was drawn for the panel survey. The SARS IRP5 data can thus be used as a check on the job flow estimates of Kerr et al. (2014). This is undertaken in Section 4 below.

3 SARS IRP5 data

In the process of collecting income taxes from workers SARS requires all firms registered for PAYE tax to issue tax certificates to all employees, the IRP5 certificate. Any firm that has employees is required by law to register for PAYE. Since the 2010 tax year, firms have been required to issue IRP5 certificates to all employees earning more than ZAR2,000 per annum. These certificates contain information on the dates during which the person was employed in the tax year, the source and amount of income earned, and the firm's PAYE reference number. This means that it is possible to construct data on worker and job flows from the IRP5 certificates for each firm that include almost all workers in tax registered firms, except those earning less than ZAR2,000 per annum.

The firm identifier in the IRP5 data is the PAYE reference number. In discussions with SARS it was noted that firms may shift employees between different payrolls, as some large firms have multiple payrolls and thus multiple PAYE reference numbers. Thus PAYE numbers are aggregated up to the company income tax number and all the analysis below is done using this firm identifier. Thus the analysis is not plant-level analysis but enterprise-level analysis of tax-paying enterprises. To create a worker identifier, an anonymized South African identity number provided in the data by SARS has been used.

3.1 Data challenges

There has been no academic analysis of the IRP5 data thus far. Using a new source of data brings both benefits and challenges. The data that is a produced is the by-product of revenue collection by SARS rather than being produced for analysis by researchers, although SARS have undertaken in-house research on the data. There are thus several data issues that have been encountered in the analysis that should be discussed.

The first is that in tax years 2010, 2011, and 2012 SARS amalgamated income source codes on the IRP5 forms for retirement income and income from employment into the 3601 code that in other years is only for employment income. This means that it is not immediately possible to distinguish between employees earning income and pensioners receiving a pension from a pension fund in these years. In internal work, SARS has estimated ratios of pensioners to working people in other years where it is possible to distinguish between them and then applied these ratios to the total number of pensioners plus employed to obtain an estimate of total employment.

It is not possible to undertake this correction when doing analysis of firm-level data. Thus pension funds have been identified by identifying firms that issued very large numbers of tax certificates that could be either for employment income or pension income in 2012 but that issued very few or no employment income tax certificates in 2013, when pension and employment income source codes were separated. Twenty-three 'firms' issuing around 1.3 million certificates in 2012 and 1.1 million certificates in 2011 (individuals can have multiple certificates or sources of pension income) fit this description and have been excluded as probable pension funds. There are approximately 500,000 certificates that are likely to be pension incomes in 2012 and 2011 which have not been identified. The inclusion of these certificates as 'jobs' in 2011 and 2012 will likely overstate the measured job and worker

reallocation between 2012 and 2013—the changeover period when pension incomes were included and then excluded from the 3601 source code.

A separate issue from the combining of pension income and income from employment source codes is that there are source codes associated with employment other than the 3601 income source code that has been used in the analysis below. SARS estimates that the 3601 source codes represent 95 per cent of all jobs. Thus roughly 5 per cent of jobs that have other income source codes are excluded from the analysis.

It was noted above that an anonymized version of the South African ID number is used as the individual identifier. This is problematic because foreigners without South African identity numbers will be excluded. There are about 310,000–350,000 certificates with no identity numbers in each year—this is less than 1 per cent of the total certificates used in each wave. Most of these are likely to be foreigners but they are not incorporated into the analysis. If firms' hiring and separation patterns, or workers' movements between firms vary between foreigners and locals the results may be affected when foreign workers are included—but any changes are likely to be small given the small fraction made up by foreign workers in the total.

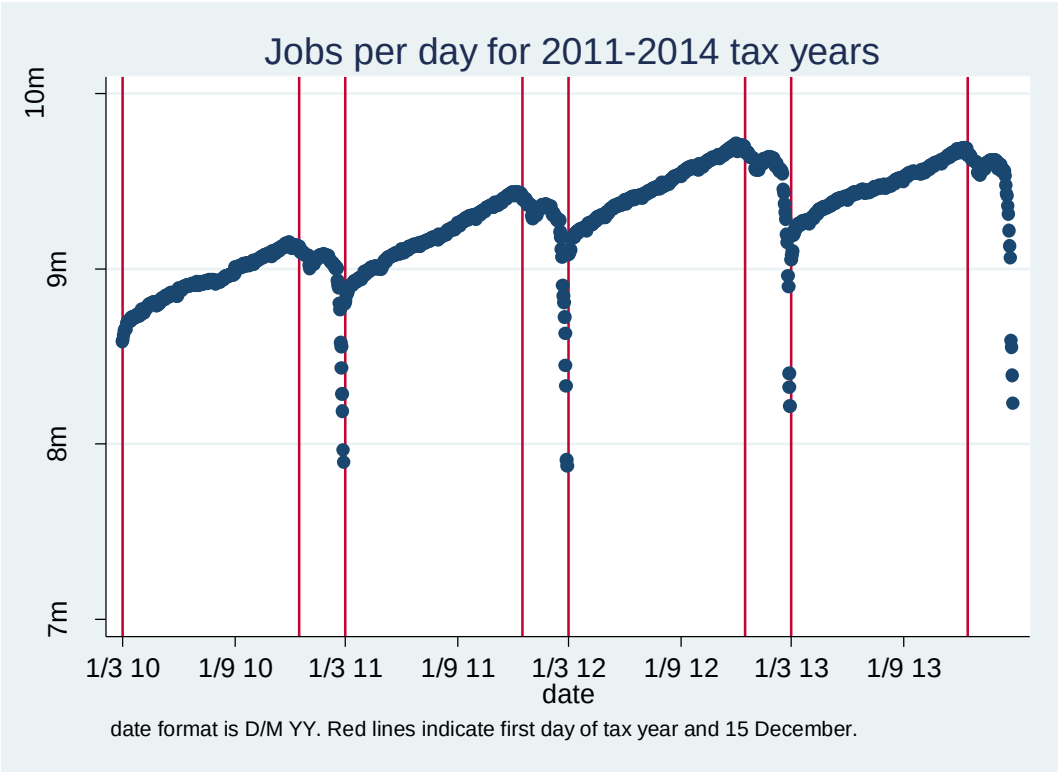
3.2 Measurement error in period employed

Each IPR5 certificate contains information on the period employed from and to in each tax year. This information can be used to determine how many jobs are being undertaken on each day of the year between 2011 and 2014. This is shown in Figure 1, which suggests a number of measurement error concerns. First is the massive drop in number of individuals employed in the last two weeks of the tax year (the last two weeks in February). The second is the large drop after the middle of December in each year. The third is that this drop seems to continue into the following tax year, so that the number of jobs only reaches the December peak in June of the following year. The last one is that, in the final tax year (ending in February 2014), the total does not surpass the previous tax year's December high, unlike in the previous three years.

Clearly there are measurement problems. It is not clear why but a number of employers (or perhaps a few large employers), do not record their employees working in the last two weeks of the tax year. The fourth problem is likely the effect of the inflation of the number of jobs in the middle two tax years due to the inclusion of certificates that are actually pension incomes but that have not been excluded (see discussion above). The drop after the middle of December is worrying and it is not clear why this occurs.

Despite these errors, the analysis below on job and worker flows requires using the information on the period employed from and to. This could be problematic given the measurement error worries. Figure 1, however, suggests that the same measurement error seems to be repeated in each year. For analysis that uses data only from the first week of the tax year this is thus less of concern. The main result of extremely high worker flows is robust to using alternative weeks within the tax year.

Figure 1: Jobs per day measured using the employed from and to variable



Source: Author's calculations from the IRP5 data.

3.3 Measuring births and deaths

The SARS data does not have direct data on births and deaths of firms, nor on mergers. A firm is treated as having died in year t if it has had workers in prior years but not in year t or any subsequent years. Similarly a firm–year observation is counted as a birth if a firm did not have any employees in years prior to year t. It is thus possible that some 'deaths' in the data may just be firms having been bought by other firms or firms that stopped producing at t but which may be continuing in subsequent years. Hethey-Maier and Schmieder (2013) find that this is a severe problem in German administrative data. In the analysis below, payrolls are aggregated to tax-paying entities which may prevent some overstatement of the extent of firm births and deaths. There is not much else that can be done about this issue other than note that it may overestimate job destruction coming from firm deaths and job creation from firm births.

4 Data preparation

The data used to estimate worker and job flows is constructed using the following method. The IRP5 data is used—all tax certificates (for incomes of more than ZAR2,000 per year) issued by all employing firms registered for PAYE between 2011 and 2014 inclusive. Only those certificates with 3601 source codes are kept—this code indicates income from employment in 2013 and 2014 and either income from employment or taxable pension income in 2011 and 2012. The 1.1 million certificates that look as if they are issued by pension funds in 2011 and 2012 are excluded, but we cannot identify approximately 500,000 other pension fund certificates which are then incorrectly counted as employment. Company income tax numbers are matched with PAYE reference numbers (payroll numbers) and the analysis is conducted on company income tax (CIT) numbers not payrolls—thus aggregating payrolls where there is more than one

payroll per CIT number. There are thus between 210,000 and 230,000 employers in each tax year.

Only individuals reporting working in the first week of the tax year (the first week in March) are kept.[4] Those with no identity numbers are dropped and, as noted above, these individuals are likely to be foreigners. Identity numbers with more than five different certificates in the first week of the tax year are also dropped, as there are a few odd-looking individuals with large numbers of certificates. This is about one-twentieth of 1 per cent of total records in each of the four years of data. Duplicates for individuals reporting multiple certificates for the same firm in the same year are also dropped, although one record for these individuals is kept.

Table 1 shows the pattern of individuals who appear in the dataset over the four years of data (to create this table, individuals who are employed in more than one firm in the same year have duplicates dropped so that there is only one observation per individual per year). There are over 12.2 million different individuals in the data. Of these, 5.4 million, or 44 per cent, appear in each of the four waves. Just over 1 million individuals appear for the first time in 2014, while approximately 900,000 individuals appeared for the first time in 2012 and then again in the following two years. The median worker appears in three of the four years of data.

Table 1: Worker panel description 2011–14

Number of workers	Per cent	Pattern
5,420,889	44.17	1111
1,072,226	8.74	...1
906,756	7.39	.111
832,546	6.78	1...
825,744	6.73	..11
678,342	5.53	111.
661,188	5.39	11..
444,800	3.62	..1.
364,573	2.97	.1..
1,066,027	8.69	(other patterns)
12,273,091	100	

Notes: In the Pattern column a '1' means the firm is present and an employer and a '.' indicates the firm was not an employer in a particular year.

Source: Author's calculations from the IRP5 data.

Measures of employment, hires, and separations are constructed in a worker-level dataset, comparing employer identifiers at year t and year t+1. Thus the turnover method is used to calculate worker flows (Huber and Smeral 2006). The data is then collapsed by firm identifier to calculate changes by firm. Thus *worker flow* analysis is conducted on a firm-level dataset created from a worker-level dataset. *Job flows* are calculated using changes in the total employment variable created when the data is collapsed by firm identifier. This data preparation results in a dataset of around 36 million certificates for 12.2 million individuals issued by approximately 290,000 employers over the four tax years 2011–14. Table 2 shows the panel of firms that are used to describe job and worker flows. Of the 290,000 firms observed at least once in the four

[4] As a robustness check the results are replicated using data from the first week in December, the period of the tax year with the maximum number of jobs. The results from this check on the data are reported below.

years around 54 per cent are observed in every year. Around 22,000 are only observed in 2014—these are assumed to be firm births. Just under 20,000 were first observed in 2012 and continued employing through 2014—these are assumed to be births in 2012.

Table 2: Firm panel description 2011–14

Number of firms	Per cent	Pattern
158,354	53.89	1111
22,179	7.55	...1
19,886	6.77	.111
18,864	6.42	1...
17,499	5.96	..11
16,776	5.71	111.
15,978	5.44	11..
4,436	1.51	.1..
4,388	1.49	1.11
15,490	5.27	(other patterns)
293,850	100	

Notes: In the Pattern column a '1' means the firm is present and an employer and a '.' indicates the firm was not an employer in a particular year.

Source: Author's calculations from the IRP5 data.

4.1 Firm-level descriptive statistics

Table 3 shows some basic data on the size of firms created using the method above. The numbers are very stable over time, suggesting that the data issues discussed above are not affecting the creation of a reliable firm-level dataset. The median firm size of six employees is slightly smaller than the median of seven calculated in the Statistics South Africa firm survey the QES by Kerr et al. (2014). This is likely due to the fact that the QES is sampled from value-added tax (VAT) registered firms and firms with turnover of less than ZAR1 million are not obliged to register for VAT. By contrast the IRP5 data should include all formal firms that have employees since all firms are required to register for PAYE. Mean firm size increases by 3 employees from 37 to 40, or around 8 per cent over the four years for which we have data.

Table 3: Firm size distribution by year using number of employees

	Year			
Percentile	2011	2012	2013	2014
1	1	1	1	1
10	1	1	1	1
25	2	3	3	3
50	6	6	6	6
75	15	15	16	16
90	40	41	42	43
99	384	391	403	410
Mean	37	38	40	40
Firm size for median worker	670	679	813	797

Source: Author's calculations from the IRP5 data.

Hsieh and Klenow (2011) discuss another important indication of how production is organized—the size of the firm where the median worker works; this is calculated by ranking workers by the size of their firm and exploring the size of firm in which workers at different percentiles in the distribution work. Table 3 shows that the median worker works in a firm of between 670 and 813 over the four years. This number is much larger than reported by Kerr et al. (2014) using the QES, who find that the median worker worked in a firm of size 140 between 2005 and 2011. The likely reason for this big difference is that the IRP5 data include a number of very large public sector organizations, the largest of which employs more than 300,000 workers. Any median weighted by the number of workers in the firm will be affected by the presence of a few very large 'firms', in a way that the regular median or mean is not. This can also be seen in the differences in the mean firm size, which go the other way and are much smaller in the IRP5 than in the QES—probably due to the extremely large number of firms with just a single employee in the IRP5 data.

5 Analysis of worker flows, job flows, and churning

5.1 Worker flows

The last column of Table 4 shows that worker flows, the sum of hires and separations, constitute around 52–4 per cent of average employment in period t and t-1 between 2012 and 2014, using the reallocation concept. Thus more than 50 per cent of those employed either left their current employer and/or arrived at their current employer in each year. This is evidence that worker flows in South Africa are very large and are a pervasive part of the way the labour market operates. This result is consistent with the conclusions of Banerjee et al. (2008: 730) using individual-level data (the Labour Force Survey—LFS—panel) for South Africa between 2002 and 2004, who note that 'Most importantly, we are struck by just how much churning there appears to be in the labour market.' By 'churning' these authors actually mean worker flow as it has been defined above, since churning cannot be calculated from worker-level data. It should be noted that the LFS data used by these authors included workers from both informal and formal sectors, while our data only includes workers working in firms registered for PAYE tax with the SARS.

International comparisons are difficult because of the different methods, time periods, and types of data that have been used to measure worker flows. But it does seem that the worker flow rates in Table 4 are large relative to many other countries for which there is data. For example in their cross-country study of worker and job flows Bassanini et al. (2010) note that 'in some countries annual job and worker reallocation are *as large as* 25% and 45%, respectively' (emphasis added) but the worker flows in Table 4 are 8 percentage points larger than the supposedly high level of 45 per cent.

Table 4: Worker flow rates by average firm size

	0–19 (%)	20–49 (%)	50–99 (%)	100–249 (%)	250–499 (%)	500–999 (%)	1,000–4,999 (%)	5,000+ (%)	All firms (%)
2012	71.7	65.7	66.5	65.8	63.6	62.3	52.6	36.9	54.1
2013	68.2	64.1	65.7	67.3	62.9	62.8	55.2	33.4	52.7
2014	67.7	62.6	64.4	66.2	61.7	63.1	52.8	35.6	52.9

Source: Author's calculations from the IRP5 data.

Bassanini et al. (2010) argue that high levels of labour reallocation as measured by worker flows are evidence of lower rigidities in the labour market. If turnover costs are high due to strong labour regulation then this will influence how firms choose to hire and fire. High turnover costs, for example a requirement for expensive legal processes before a worker can be fired, are likely to reduce worker flows. Thus the estimates of worker flows shed some light on the extent to which South African firms are constrained by labour legislation, which is a common concern raised by the business community and suggests that rigidities on hiring and firing may not be as much of a concern as had previously been thought (Go et al. 2009). However, Martin and Scarpetta's view that only parts of the labour market may be highly regulated and that this would lead to inefficiently high worker flows in the unregulated parts is also a possibility.

As found in much previous research (Burgess et al. 2000; Huber and Smeral 2006), Table 4 shows that flow rates are declining in average firm size in each of the three years of data that these flow rates can be measured. Worker flows are just below 70 per cent in firms of fewer than 20 employees but only around 35 per cent for firms with more than 5,000 workers.

Table 5 shows worker flows by firm median earnings quintile. To calculate the quintiles required the use of the date employed from and to on the tax certificate (from which a months employed variable was created) and the total earnings for the tax year reported for the 3601 source code—which is income from employment. From these variables a monthly income variable for each individual was created. Monthly individual incomes of the top 1 per cent of earners and the bottom 1 per cent of earners were set to missing as a result of measurement error concerns.[5] A median within-firm monthly wage was then created for each firm in each year. For the analysis in Table 5 firms were then divided into quintiles by the median wage in the firm. To make this constant over the four years of data the mode of the four median wage quintiles for each firm was used. The monthly income variable will suffer from the same measurement issues discussed above and shown in Figure 1, as it is calculated from the period employed from and to. This means that firms that do not report their employees working in the last two weeks in the tax year (see Figure 1) would have slightly overestimated higher monthly earnings compared to firms that do not seem to have made this error. It is not clear what effect this will have on the results.

Table 5: Worker flow rates by firm median monthly earnings quintile

	Quintile 1 (%)	Quintile 2 (%)	Quintile 3 (%)	Quintile 4 (%)	Quintile 5 (%)	All (%)
2012	79.9	67.0	57.1	41.3	33.1	54.1
2013	80.3	65.1	58.2	40.4	29.6	52.7
2014	82.8	65.4	56.3	40.0	29.6	52.9

Source: Author's calculations from the IRP5 data.

There is substantial variation in worker flow rates across industries, as shown in column 2 of Table 6. The largest worker flow rates are 93 per cent in 'Other manufacturing', 79 per cent in 'Personal and household services' and 72 per cent in 'Catering and accommodation'. The smallest flow rates are 20 per cent per year in 'Public administration', 35 per cent in 'Mining and quarrying' and 37 per cent in 'Electricity, gas, and water' (which also contains a number of state-owned enterprises). These industrial classifications are obtained from firms self-reporting their main activity on their EMP501 returns. SARS then classifies these reports into 34 industries, although this is not the standard industrial classification (SIC) system, nor is it likely to be as

[5] There is a non-negligible fraction of individuals that firms reported earning R1 in the year. It is not clear why this occurs.

reliable as the Statistics South Africa classification, for which Statistics South Africa spends time contacting firms to obtain improved classifications. The largest industries by average number of employees are 'Finance, insurance, real estate+ business services', 'Agencies and other services', and 'Public administration'.

Table 6: Industry-level statistics

Self-reported industry classification	Average no. of jobs	WFR average	JRR average	Churning average
Agriculture, forestry, and fishing	391,134	0.68	0.32	0.36
Mining and quarrying	321,324	0.35	0.15	0.20
Food, drink, and tobacco	103,415	0.57	0.27	0.30
Textiles	24,188	0.48	0.26	0.22
Clothing and footwear	84,343	0.56	0.13	0.43
Leather, leather goods + fur	4,173	0.45	0.26	0.19
Wood, wood products, and furniture	60,945	0.51	0.23	0.27
Paper, printing, and publishing	59,169	0.48	0.22	0.26
Chemicals, rubber, and plastic products	61,800	0.51	0.25	0.26
Coal and petroleum products	13,129	0.59	0.31	0.27
Bricks, ceramics, glass, cement	37,552	0.51	0.23	0.28
Metal	67,162	0.41	0.20	0.21
Metal products	66,137	0.52	0.28	0.24
Machinery and related items	88,998	0.50	0.24	0.25
Vehicles, parts, and accessories	69,850	0.44	0.19	0.25
Transport equipment	11,757	0.57	0.31	0.26
Scientific, optical, and similar equipment	4,417	0.59	0.30	0.29
Other manufacturing industries	96,132	0.93	0.35	0.58
Electricity, gas, and water	63,295	0.37	0.24	0.13
Construction	371,645	0.74	0.36	0.38
Wholesale trade	314,824	0.57	0.22	0.35
Retail trade	412,090	0.61	0.26	0.35
Catering and accommodation	155,703	0.72	0.31	0.41
Transport, storage, and communication	300,513	0.48	0.22	0.26
Finance, insurance, real estate + business services	2,194,885	0.62	0.29	0.33
Public administration	1,179,764	0.20	0.07	0.13
Educational services	246,724	0.41	0.13	0.28
Research and scientific institutes	10,659	0.63	0.26	0.37
Medical, dental, veterinary services	104,299	0.53	0.24	0.29
Other social + related community services	91,039	0.51	0.24	0.27
Recreational and cultural services	35,951	0.64	0.31	0.34
Personal and household services	55,877	0.79	0.30	0.49
Specialized repair services	49,320	0.61	0.31	0.29
Agencies and other services	1,329,671	0.58	0.22	0.36

Notes: WFR—worker flow rate; JRR—job reallocation rate.

Source: Author's calculations from the IRP5 data.

5.1.1 Individual heterogeneity in worker flows

Burgess et al. (2000) make the point that high levels of worker flows and churning do not mean most workers are churned. Worker experiences will depend on the heterogeneity of flows across workers. Table 7 shows the fraction of employees in surviving firms that are still employed in these firms one, two, and three years later (for the SARS IRP5 data) and for longer periods for the US. For both countries there are clearly some workers experiencing high levels of stability—with 78 per cent of workers still employed by the same firms after one year, 65 per cent after two years, and 56 per cent after three years in the South African IRP5 data, and similar numbers in the US data presented by Burgess et al. (2000).[6]

Table 7: Percentage of workers in the same firm in subsequent years

2011	SARS IRP5 data	1985	State of Maryland	
2012	77.9		Non-manuf.	Manuf.
2013	65.4	1986	72.1	75.5
2014	56.2	1987	62.2	66.6
		1988	59.8	64.3
		1990	47.3	42.5
		1994	42.1	31.9

Notes: Both sets of results are for surviving firms only.

Source: Author's calculations from the IRP5 data and Burgess et al. (2000) for the Maryland data.

5.2 Job flows

The SARS IRP5 data can also be used to measure job reallocation. Table 8 shows job creation and destruction rates by firm size and for all firms in the final column. Job creation is between 11 and 14 per cent in the three years of data, while job destruction is slightly lower—around 10 per cent. This means that job reallocation—the sum of the absolute value of job creation and destruction—is between 21 and 24 per cent over the three-year period between 2012 and 2014 that we can estimate it for. These numbers include public sector entities, which employ about one-eighth of all workers in the IRP5 data on average over the years for which we have data.

Like worker flows, job reallocation is higher in smaller firms. Table 8 shows that job reallocation rates are about 45 per cent in firms with fewer than 20 employees but only 10.5 per cent in the largest firms. It is also possible to explore job creation and destruction rates in firms of differing wage rates. Table 9 shows job creation and destruction by wage quintile, calculated in the same manner as described above. Job reallocation rates (the sum of job creation and destruction rates) are monotonically declining in wage rates so that the firms with the lowest wage rates have the highest levels of churning. Because this analysis is by firm quintile the number of firms in each quintile is the same but there is substantially more employment in the firms in the upper quintiles since higher-paying firms are generally larger. These results are not weighted to reflect these differences in employment.

Job reallocation rates by the 34 main industry categories available in the IRP5 data are shown in the third (JRR) column of Table 6. Job reallocation rates are lowest in 'Public administration' (6.8 per cent on average per year), 'Educational services' (13 per cent) and 'Mining and quarrying' (15

[6] These numbers are not comparable to the worker flow figures above, because they only include workers in surviving firms.

per cent) and highest in 'Construction' (35 per cent), 'Agriculture, forestry, and fishing' (34 per cent) and 'Catering and accommodation' (31 per cent).

Table 8: Job reallocation by firm size

Job destruction	Firm size 0–19 (%)	20–49 (%)	50–99 (%)	100–249 (%)	250–499 (%)	500–999 (%)	1,000–4,999 (%)	5,000+ (%)	All firms (%)
2012	-23.4	-16.4	-15.4	-13.1	-12.1	-11.4	-7.3	-4.4	-10.7
2013	-21.3	-14.3	-12.5	-12.0	-10.8	-8.5	-10.4	-3.0	-9.6
2014	-20.7	-14.5	-12.8	-13.7	-11.6	-11.5	-8.9	-5.7	-10.8
Job creation									
2012	23.9	18.9	18.0	16.5	15.8	12.7	13.2	8.5	14.1
2013	22.6	18.9	18.6	17.8	14.6	14.6	10.3	6.1	12.8
2014	22.9	17.2	17.2	15.8	13.5	12.8	9.3	3.8	11.4
Job reallocation									
2012	47.3	35.3	33.4	29.7	27.9	24.1	20.5	12.9	24.8
2013	43.9	33.2	31.1	29.8	25.5	23.1	20.7	9.1	22.4
2014	43.6	31.6	30.0	29.5	25.1	24.2	18.1	9.5	22.1

Source: Author's calculations from the IRP5 data.

Table 9: Job reallocation rates by firm median earnings quintile

	Quintile 1 (%)	Quintile 2 (%)	Quintile 3 (%)	Quintile 4 (%)	Quintile 5 (%)	All (%)
2012	37.6	29.8	26.9	17.9	15.2	24.8
2013	35.7	27.8	26.2	15.3	12.0	22.4
2014	35.2	27.6	23.9	16.6	11.8	22.1

Source: Author's calculations from the IRP5 data.

Given that this is the first time the IRP5 data have been analysed, it is of interest to compare the results to previous studies using other sources of data. Kerr et al. (2014) found that yearly job reallocation rates were on average 20 per cent using Statistics South Africa QES data from 2005–11, meaning that the estimates using SARS IRP5 data from 2011–14 are 3 percentage points higher on average. But the difference between the two sets of results is actually larger than this because the QES study excluded government departments, agriculture and mining, while these sectors are included in the above analysis of the SARS IRP5 data. When excluding firms in these three sectors, the average job reallocation rate in the IRP5 data increases to an average of 26.8 per cent, nearly 7 percentage points and nearly 30 per cent higher than the Kerr et al. (2014) estimates, mainly due to the extremely low level of job reallocation in the public sector. Kerr et al. (2014) did note that the QES was likely to underestimate job reallocation because births were not well captured, and because small firm job creation and destruction was missed every time a new panel was sampled—these are likely to be the causes of the differences in the two sets of results. The QES data also excluded smaller firms because only firms registered for VAT appeared in the QES sample. Since smaller firms generally have higher amounts of job reallocation this may also explain the higher IRP5 number.

Table 10 shows the contribution of births and deaths to job creation and destruction in the IRP5 data. Deaths contribute about 34 per cent of total job destruction while births contribute about 21 per cent of total job creation. Kerr et al. (2014) found that deaths contributed 26 per cent of total job destruction on average between 2005 and 2011 in the QES. It was noted above that some of the 'deaths' in the SARS IRP5 data may actually be mergers or temporary shutdowns. The QES data used by Kerr et al. (2014) identified both mergers and deaths directly, whereas this has to be inferred from the IRP5 SARS data, suggesting that the numbers from the QES data on deaths are more reliable. If the higher figure for death-related job destruction in the IRP5 is due solely to the misclassification of mergers as deaths then, assuming the QES number is true, job destruction would be 8 per cent lower and job reallocation would be slightly less than 1 percentage point less than estimated in Table 8. This means overestimation of deaths is not likely to be the explanation for the substantially higher estimate of job reallocation using the SARS data—this is more likely the result of the weaknesses of the QES data for measuring job reallocation, as discussed above.

Table 10: Job destruction and creation from firm deaths and births

	Firm size								
% job destruction from deaths	0–19 (%)	20–49 (%)	50–99 (%)	100–249 (%)	250–499 (%)	500–999 (%)	1,000–4,999 (%)	5,000+ (%)	All firms (%)
2012	42.1	29.4	28.8	25.5	34.7	38.6	33.7	0.0	29.5
2013	44.2	34.0	32.5	32.4	31.9	27.1	27.7	31.3	34.4
2014	56.2	43.5	39.4	35.8	36.1	39.8	30.9	13.3	37.2
% job creation from births									
2012	48.2	34.3	28.5	23.3	19.6	17.0	11.0	27.6	28.5
2013	35.6	21.9	16.7	16.0	8.4	6.8	10.8	3.1	16.7
2014	33.4	23.8	19.7	19.0	14.9	7.7	10.2	0.0	18.6

Source: Author's calculations from the IRP5 data.

It is also possible to explore the contributions of firm births and deaths of firms of different sizes. As in the results of Kerr et al. (2014), the contributions to job reallocation are much higher for smaller firms. Table 10 shows that deaths contribute around 48 per cent of job destruction in firms of fewer than 20 employees while births contribute around 39 per cent of job creation in these firms. There is not a monotonic decline in the contribution of births and deaths to job creation and destruction however. This is a similar finding to that of Kerr et al. (2014) for firm deaths. Firm births were not well measured in the QES so comparisons on this measure are not useful.

5.3 The robustness of worker and job flow estimates

It was noted above that there is some measurement error in the period employed from and to variables. One robustness check that has been undertaken is repeating the worker flow and job reallocation analysis on an identical dataset except that the period of employment used was the first week in December rather than the first week in March (which is the first week of the tax year). The first week in December was the period of maximum measured employment, shown in Figure 1. The analysis using the December data produced estimates of worker and job flows and patterns by industry, wage quintile, and firm size that are not substantially different to the March results. Worker flows are around 3 percentage points higher and the job reallocation rates are

less than 1 percentage point higher using the data from the first week in December. This is thus one test that suggests that measurement error in period of employment is not the cause of the high worker flows discussed above.

5.4 Churning

Churning flows are the difference between worker flows and job flows and measure the amount of worker reallocation occurring over and above that required due to the job reallocation undertaken by firms. Table 11 shows churning flow rates in South Africa. One way to measure the importance of churning is to calculate the proportion of churning in the overall worker flow rate (Burgess et al. 2000). Table 12 implies that between 54 per cent and 58 per cent of worker flows are churning flows in each year between 2011 and 2014 in South Africa. This ratio method gives a minimum of 34 per cent in the smallest firms (Table 12) and a maximum of 73 per cent in the largest firms. The other method used by Burgess et al. (2000) is to take the mean of the ratio of worker flows to churning flows across all employers. This gives a slightly lower result of 46 per cent across all years and employers in the IRP5 data.

Table 11: Churning flow rate by firm size

	0–19 (%)	20–49 (%)	50–99 (%)	100–249 (%)	250–499 (%)	500–999 (%)	1,000–4,999 (%)	5,000+ (%)	All firms (%)
2012	24.44	30.41	33.07	36.11	35.72	38.20	32.10	23.96	29.33
2013	24.32	30.92	34.62	37.46	37.43	39.71	34.55	24.28	30.26
2014	24.09	30.99	34.40	36.66	36.61	38.86	34.65	26.08	30.78

Source: Author's calculations from the IRP5 data.

Table 12: Churning as a percentage of worker flows by firm size

	0–19 (%)	20–49 (%)	50–99 (%)	100–249 (%)	250–499 (%)	500–999 (%)	1,000–4,999 (%)	5,000+ (%)	All firms (%)
2012	34.1	46.3	49.7	54.9	56.2	61.3	61.0	64.9	54.2
2013	35.7	48.2	52.7	55.7	59.5	63.2	62.6	72.7	57.4
2014	35.6	49.5	53.4	55.4	59.3	61.6	65.6	73.3	58.2

Source: Author's calculations from the IRP5 data.

The importance of churning flows in worker flows in South Africa is less than found by Burgess et al. (2000) for the US state of Maryland, who estimated a mean of 70 per cent in non-manufacturing and 62 per cent in manufacturing, taking the ratio across all employers. It may indicate more churning in the US than in South Africa. This difference may be also partly because, for their work, Burgess et al. (2000) drop firms with fewer than five workers, which are likely to have higher than average churning flows. When dropping firms with an average of fewer than five workers the ratio of churning to worker flows across all firms increases from 46 per cent to 53 per cent in the SARS IRP5 data. It should be noted that this method weights all firms equally—whether a firm has 1 or 10,000 employees—so differences in the firm size distribution between South Africa and the US may also play a role. It should be pointed out that much of the following analysis of churning is conducted at the firm level, without weighting by employment, following the international literature.

The last column of Table 6 shows that churning flows do differ substantially by industry. They are as low as 13 per cent of employment in 'Public administration' and 'Electricity, gas, and water', and as high as 58 per cent in 'Other manufacturing', 49 per cent in 'Personal and

household services' and 43 per cent in 'Clothing and footwear' (all three of which have fewer than 100,000 jobs on average over the period 2011–14).

5.5 Heterogeneity in worker and job flows

'The pervasiveness of heterogeneity and diversity in economic life' is a key insight of recent empirical research in economics (Heckman 2001: 674). Understanding the nature of heterogeneity in both firms and workers has been an important goal of labour economists in recent times (Haltiwanger et al. 2007). Davis et al. (1996) highlighted the great heterogeneity of job reallocation rates in US manufacturing firms, even within narrowly defined industries. The authors showed that industry dummies explained only a very small fraction of firm-level job flow rates and argued that industry-level analysis may not be helpful because it did not take into account the heterogeneity of outcomes within industries. Burgess et al. (2000) strengthened this argument by documenting the unimportance of industry effects in explaining churning and worker flow rates. This work is important because it suggests that any industry-level analysis misses the large within-industry heterogeneity.

This argument can be explored for South Africa using the IRP5 data. Table 13 shows the R^2 from regressions of combinations of employer or four-digit industry fixed effects and time effects on firm-level job reallocation rates, job flow rates, worker flow rates, and churning rates using the IRP5 data. Rows 3 and 4 of the table show that industry, or industry plus time effects explain no more than 2.3 per cent of the variation in any of these firm-level outcomes, and show that in South Africa, as in many countries, there is substantial heterogeneity in firm-level job and worker flow outcomes within even narrowly defined industries. Adding in employer effects does not substantially increase the explanatory power of the regression of job flow rates, even though there are only three observations per firm, confirming a result from Burgess et al. (2000) that job flows have a substantial idiosyncratic component. Adding in employer fixed effects, however, does increase the R^2 of worker flow and churning rates, implying that there are important differences across firms. This evidence can be interpreted as supporting the argument by Burgess et al. (2000) that a variety of equilibrium human resource policies are pursued by different firms even within narrowly defined industries.

Table 13: The explanatory power of employer, industry, and time effects

	JRR	WFR	JFR	CR
Employer	0.354	0.409	0.03	0.32
Employer, time	0.355	0.409	0.086	0.32
Industry	0.014	0.023	0.001	0.013
Industry, time	0.022	0.03	0.011	0.014

Notes: The table shows the R^2 from regressions of the column heading variable on the row characteristics (a combination of either employer or industry fixed effects and time dummies). JRR—job reallocation rate; WFR—worker flow rate; JFR—job flow rate; CR—churning rate.

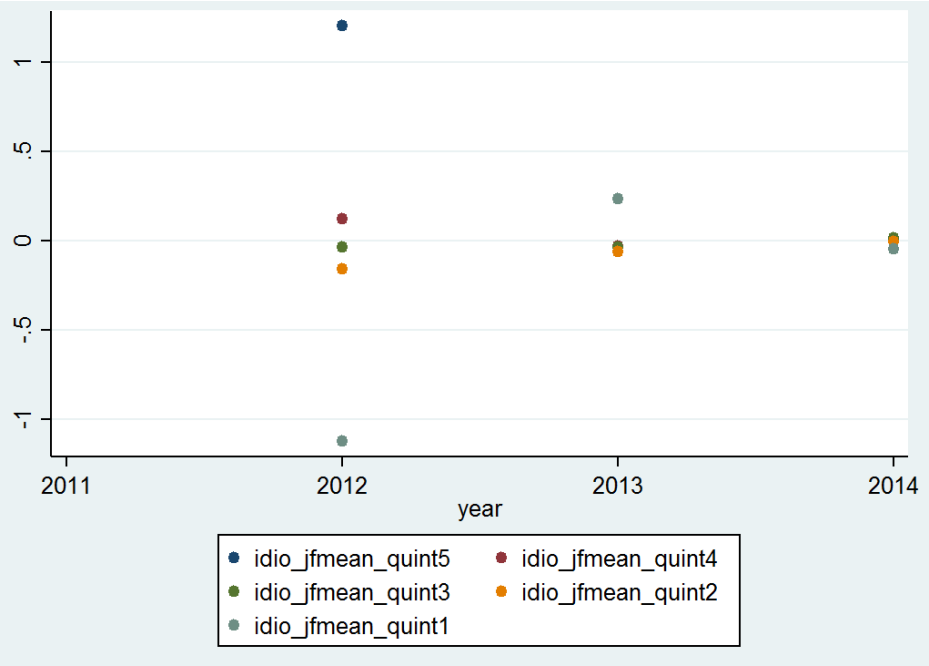
Source: Author's calculations from the IRP5 data.

Further evidence in support of this argument can be illustrated using the idiosyncratic churning rate, the churning rate of firm *i* minus the average of the four-digit industry of the firm at time *t*.[7] Following Burgess et al. (2000) firms are divided into quintiles based on idiosyncratic churning rates from 2012. Figure 2 shows a plot of the average idiosyncratic churning rate by initial 2012

[7] In addition to the 34 industries used in Table 6 the IRP5 data also has a four-digit*?* industry code derived from a self-classification provided by each that is used in this analysis.

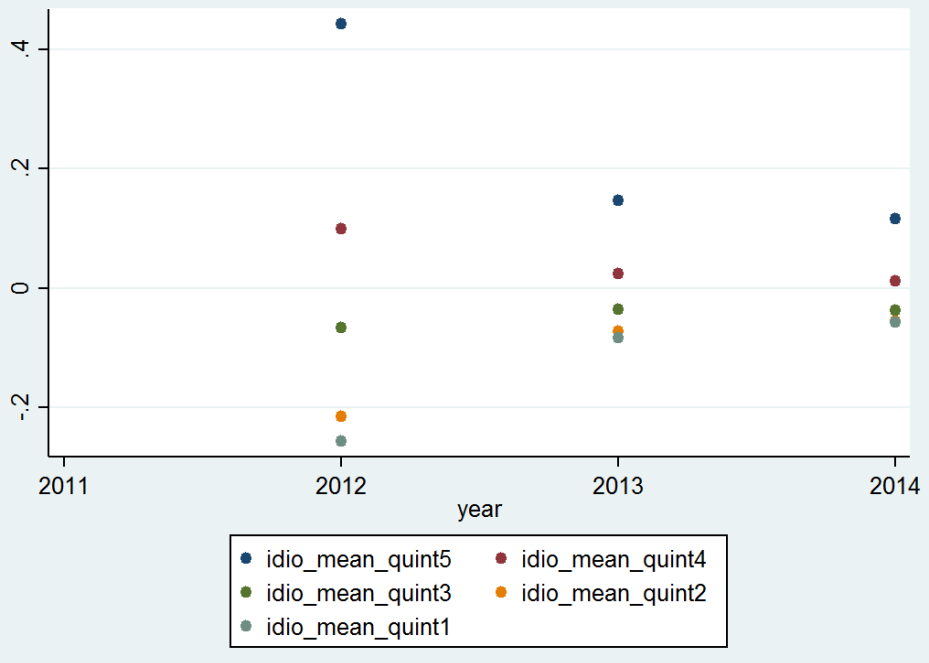
quintile in 2013 and 2014. While not as strong as the results in Burgess et al. (2000) (and being for a much shorter time period), they show that there is some persistence in idiosyncratic churning rates—those firms in the top quintile of idiosyncratic churning rates in 2012 have the highest average churning compared to the firms in the other quintiles in 2013 and 2014. Again this suggests that different human resource policies may exist in different firms. Figure 3 shows, though, that there is no such persistence in idiosyncratic job flow rates but rather strong regression to the mean.

Figure 2: Job flow rates by 2012 idiosyncratic job flow rate quintile



Source: Author's calculations from the IRP5 data.

Figure 3: Churning rates by 2012 idiosyncratic churning rate quintile



Source: Author's calculations from the IRP5 data.

One possible choice of human resource policy that firms have is to trade off wages and churning rates. Table 14 shows the results of a regression of firm-level churning rates on firm characteristics and shows that firms that pay lower wages do have higher churning rates. This is true in regressions with current median wage in the firm or previous year's median wage as explanatory variables. The interpretation of the coefficient on the idiosyncratic current median wage in the firm is that a ZAR1,000 increase in the monthly idiosyncratic wage of a firm would decrease the churning rate by 0.3 percentage points. The coefficient on the log of employment variable implies that a 1 per cent increase in employment is associated with a .036 percentage point increase in the idiosyncratic churning rate. Both results are statistically significant at all conventional levels.

The results from this section confirm that churning and worker flows are not random. Some South African firms have very high churning and worker flow rates and others have much lower rates. Churning rates exhibit some persistence over time within firms, unlike job flows which exhibit strong regression to the mean. It has also been shown that firms that pay lower wages have higher churning rates, suggesting that in equilibrium there can be both high-wage/low-churn and low-wage/high-churn policies implemented by different firms within the same industries.

Table 14: The determinants of idiosyncratic churning

Variables	(1)	(2)
Idiosyncratic monthly earn median (*10,000)	-0.0309***	
	(0.000876)	
Log (employment)	0.0363***	0.0342***
	(0.000397)	(0.000405)
2013.year	-0.000724	-0.0186***
	(0.00110)	(0.00113)
2014.year	7.90e-05	-0.0197***
	(0.00110)	(0.00113)
Lag idiosyncratic monthly earn median (*10,000)		-0.0397***
		(0.000941)
Constant	-0.0808***	-0.0501***
	(0.00134)	(0.00140)
Observations	366,480	350,039
R^2	0.026	0.026

Notes: Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Both regressions are of the firm's idiosyncratic churning rate on other firm characteristics and time dummies. Column (1) includes current median earnings and column (2) includes lagged median earnings.

Source: Author's calculations from the IRP5 data.

6 Conclusions

The SARS IRP5 data can be used to measure the extent of job flows, worker flows, and churning in South Africa and shows that South Africa has substantial levels of worker flows—around 53 per cent per annum. These flows are higher than would be expected given the extremely high levels of unemployment and the reputation the labour market has for being rigid. They do,

however, confirm the finding made by Banerjee et al. (2008) using household survey panel data that there are substantial worker flows in the data.

Worker flows are found to be highest in small firms and decrease monotonically across average firm size categories. They are also highest in firms that pay the lowest wages, again declining monotonically with the median wage paid by the firm. Worker flows (and job flows) will be overestimated if firm mergers result in exaggerated levels of firm deaths or if firms shift employees across different PAYE numbers within the same tax-paying entity. To counter the second issue firm payroll identifiers are aggregated to the tax number and all analysis used this number as the firm identifier.

The IRP5 data can also be used to measure job creation and destruction, and job reallocation rates of 23 per cent were found when using the full dataset. This increased to 26.8 per cent when excluding PAYE registered entities classified as mining, agriculture, and public administration, which is substantially higher than the average job reallocation rate of 20 per cent found by Kerr et al. (2014) using private sector firms outside of mining and agriculture in the QES survey from 2005–11. Exaggeration of firm deaths in the IRP5 data can account for at most around 1 percentage point of the difference if it is assumed that the difference in the contribution of firm deaths between the QES and the IRP5 is all due to measurement error and that these extra deaths resulted in no job destruction. This suggests that the QES limitations discussed by Kerr et al. (2014) had important consequences for the measurement of job reallocation rates in that data.

Both worker flows and job flows are found to vary substantially by industry. They are lowest in public administration and gas, electricity, and water (which has a number of state-owned enterprises) and mining, and high in construction, agriculture, and a few smaller sectors. Churning, the difference between worker flows and job flows, was also found to be high, although not as important in its contribution to worker flows as in the US. Again, churning was lowest in the public sector and highest in a few very small sectors, followed by catering and construction.

The data used to create the employment variable during the first two weeks of each tax year is subject to measurement error, although the patterns it generates are repeated in each year of data used. The main results of the paper are robust to using an alternative dataset created using the first two weeks in December as the employment period to calculate job and worker flows over time. The similarity of the firm size distribution to the QES (with some explainable differences) is another reason to believe that this measurement error in period employed is not the cause of our main results.

The paper replicated some key results from the literature on the heterogeneity of worker and job flow outcomes within narrowly defined industries. Worker and churning flows are well explained by employer fixed effects, whereas job flows are not and this, as well as the regression results, suggest that firms trade off churning rates and wages.

South Africa has a vast unemployment rate. Many researchers have blamed this outcome on a rigid labour market. The results of this paper suggest that in terms of quantity adjustment the labour market is unlikely to be rigid. Future research should examine the adjustment of wages when there are output shocks.

References

Banerjee, A., S. Galiani, J. Levinsohn, Z. McLaren, and I. Woolard (2008). 'Why Has Unemployment Risen in the New South Africa?'. *Economics of Transition*, 16(4): 715–40.

Bassanini, A., A. Garnero, P. Marianna, and S. Martin (2010). 'Institutional Determinants of Worker Flows: A Cross-country/Cross-industry Approach'. OECD Social, Employment and Migration Papers 107. Paris: OECD Publishing.

Burgess, S., J. Lane, and D. Stevens (2000). 'Job Flows, Worker Flows, and Churning'. *Journal of Labor Economics*, 18(3): 473–502.

Criscuolo, C., P.N. Gal, and C. Menon (2014). 'The Dynamics of Employment Growth: New Evidence from 18 Countries'. OECD Science, Technology and Industry Policy Papers 14. Paris: OECD Publishing.

Dadam, V., and N. Viegi (2015). 'Labour Market and Monetary Policy in South Africa'. Working Paper 551. Cape Town: Economic Research Southern Africa.

Davis, S.J., and J. Haltiwanger (1998). 'Measuring Gross Worker and Job Flows'. In J. Haltiwanger, M.E. Manser, and R. Topel (eds), *Labor Statistics Measurement Issues*. Chicago: University of Chicago Press.

Davis, S.J., J. Haltiwanger, and S. Schuh (1996). *Job Creation and Destruction*. Cambridge, MA: MIT Press.

Go, D.S., M. Kearney, V. Korman, S. Robinson, and K. Thierfelder (2009). 'Wage Subsidy and Labor Market Flexibility in South Africa'. Policy Research Working Paper Series 4871. Washington, DC: World Bank.

Haltiwanger, J.C., J.I. Lane, and J.R. Spletzer (2007). 'Wages, Productivity, and the Dynamic Interaction of Businesses and Workers'. *Labour Economics*, 14(3): 575–602.

Haltiwanger, J., S. Scarpetta, and H. Schweiger (2008). 'Assessing Job Flows Cross Countries: The Role of Industry, Firm Size and Regulations'. NBER Working Paper 13920. Cambridge, MA: NBER.

Hamermesh, D.S., W.H.J. Hassink, and J.C. Van Ours (1996). 'Job Turnover and Labor Turnover: A Taxonomy of Employment Dynamics'. *Annales d'Économie et de Statistique*, 41–42: 21–40.

Heckman, J.J. (2001). 'Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture'. *Journal of Political Economy*, 109(4): 673–748.

Hethey-Maier, T., and J. Schmieder (2013). 'Does the Use of Worker Flows Improve the Analysis of Establishment Turnover: Evidence from German Administrative Data'. *Schmollers Jahrbuch—Journal of Applied Social Science Studies*, 133(4): 477–510.

Hsieh, C.T., and P. Klenow (2011). 'The Life Cycle of Plants in India and Mexico'. Chicago Booth Research Paper 11–38. Available at: http://ssrn.com/abstract=1935473 (accessed April 2016).

Huber, P., and K. Smeral (2006). 'Measuring Worker Flows'. *Applied Economics*, 38: 1689–95.

Kerr, A., M. Wittenberg, and J. Arrow (2014). 'Job Creation and Destruction in South Africa'. *South African Journal of Economics*, 82(1): 1–18.

Kingdon, G., J. Sandefur, and F. Teal (2006). 'Labour Market Flexibility, Wages and Incomes in Sub-Saharan Africa in the 1990s'. *African Development Review*, 18(3): 392–427.

Magruder, J. (2012). 'High Unemployment Yet Few Small Firms: The Role of Centralized Bargaining in South Africa'. *American Economic Journal: Applied Economics*, 4(3): 138–66.

Martin, J. and S. Scarpetta (2012). 'Setting It Right: Employment Protection, Labour Reallocation and Productivity'. *De Economist*, 160(2): 89–116.

Mortensen, D.T., and C.A. Pissarides (1994). 'Job Creation and Job Destruction in the Theory of Unemployment'. *Review of Economic Studies*, 61(3): 397–415.

OECD (2009). *OECD Employment Outlook 2009: Tackling the Jobs Crisis*. Paris: OECD Publishing.